**Amendments to the Specification**


Please replace paragraph 28 on page 8, with the following rewritten paragraph:


[0028] In block 214, the search engine 130 determines whether ~~duplication~~ duplicate detection is to be performed. If so, processing continues to block 216, otherwise, processing continues to block 218. In block 216, the duplicate detection component 138 detects duplicate documents in different redirect chains and merges the redirect chains. In certain implementations, the duplicate detection component 138 uses a content-based duplicate detection technique that uses information about the documents in the redirect chain (i.e., the equivalence class) in the "union find" data structure. In certain embodiments, two documents may be considered to be duplicates if they are similar (e.g., more than some percentage (e.g., 90% ) of their content is the same). For example, if a first and second document are considered equivalent by content, and if a third document redirects to the first document and a fourth document redirects to the second document, the redirect component 136 concludes that the first, second, third, and fourth documents are equivalent. Additionally, if a redirect chain has a first document and another redirect chain has a second document, and if the first document and second document are duplicates based on content, the redirect chains containing the first and second documents are merged to form one redirect chain.


Please replace paragraph 33 on page 9, with the following rewritten paragraph:


[0033] In certain implementations, paths of each document in the equivalence class may also be propagated to the selected representative for global analysis, which is described further in United States Patent Application No. [[xx/xxx,xxx]] 10/764,772, entitled "A PIPELINED ARCHITECTURE FOR GLOBAL ANALYSIS AND INDEX BUILDING," by Marcus F. Fontoura et al., Docket No. SVL920030120US1, filed on the same date herewith, and which is incorporated by reference herein in its entirety.